Blockchain application in remote condition monitoring

Alzahrani, Rahma; Herko, Simon; Easton, John

DOI:
10.1109/BigData50022.2020.9377895

Blockchain Application in Remote Condition Monitoring

1st Rahma A Alzahrani
School of Engineering (EESE)
University of Birmingham
Birmingham, UK
raa926@bham.ac.uk

2nd Simon J Herko
Iconic Blockchain
TravelSpirit Foundation
Atherton, UK
simon.herko@travelspirit.io

3rd John M Easton
School of Engineering (EESE)
University of Birmingham
Birmingham, UK
J.M.Easton@bham.ac.uk

***Abstract*—Through advanced sensor technologies, satellite-based authentication, and high bandwidth data networks, Remote Condition Monitoring (RCM) systems are now an essential 'Internet of Things' (IoT) resource for efficient operation of railway infrastructure. However, the full potential of this big data has yet to be realized. Data is currently collected and used in siloes, with limited visibility of all possible datasets for exploitation. The RSSB on behalf of the UK Rail Industry established a cross-industry research program, T1010, to build stronger cooperation between stakeholders in sharing RCM data. This research builds upon T1010, to explore the use of blockchain and smart contracts to automate, in an auditable and tamper-proof way, the commercial agreements and payment processes for data trading. By removing the limitations of paper-based agreements, our goal is to enable innovation in shared business processes and an IoT data marketplace. Building on existing smart contract-based schemes for trading and sharing IoT data over blockchain networks, this research identifies novel ways to enforce agreements and ensure fair cost attribution between parties, without a Trusted Third Party. The initial design of a blockchain-based framework is presented, oriented around the data provider, consumer, and smart contracts. Blockchain-hosted data access agreement and accounting models are specified in detail. The processors in the efficient permissioned blockchain platforms Hyperledger Fabric, Sawtooth, and Iroha have been analyzed for their suitability for implementation. We then outline our future work to evaluate and validate two industrial use cases: monitoring systems for unattended overhead line equipment and axle bearings.**

***Index Terms*—Big data, Blockchain, Remote Condition Monitoring, Cost attribution, Process automation.**

I. Introduction

The pursuit towards higher quality services in the railway industry is a continuous process. Many technologies have been integrated to support a transformation from a mechanic and electronic-based environment to a more informatics-based one. Remote condition Monitoring (RCM) is one prevalent technology that is used to enhance the maintainability, accessibility, safety, and reliability level of the whole railway system. By applying this technology, detecting and diagnosing persistent and imminent faults will be possible. Thereby, preventive maintenance could be achieved, which helps in avoiding breakdowns of the system and costly failures and delays. To this end, advanced computing and sensing have become a core issue to fulfill the increasing demand for monitoring the health of railway assets throughout the day and enabling maintenance to be done in a timely fashion. Therefore, more sensors and smart devices are integrated continuously and are generating data on a massive scale. In general, activities in railway RCM can be classified into four main categories (quadrants) based on the monitoring sensor location and which asset they are monitoring, namely train monitoring train, infrastructure monitoring infrastructure, train monitoring Infrastructure, and infrastructure monitoring train [1]. Smart sensors mounted on assets belong to one stakeholder but used to monitor assets belonging to another will be in the quadrant train monitoring infrastructure and vice versa. An example is sensors mounted on fixed infrastructure that are used for checking wheel flats on rolling stock [2]. In such a situation, a stakeholder accruing the business benefit from using the system may not be the one who pays the cost of installing and operating the sensing hardware. As a result, adopting technologies that have net business benefits to the whole system in the rail industry may not be a main goal for most stakeholders. This issue is encountered in many industrial systems and is expected to be exacerbated by the emergence of the availability of Internet of Things (IoT) technology.

To solve this issue, it is essential to build stronger co-operation between rail industry stakeholders with equipment either mounted on trains to monitor the infrastructure or on the infrastructure to monitor trains. This cooperation will lead to a full exploitation of RCM in the rail industry by sharing of RCM data across the rail industry. Therefore, the RSSB on behalf of the Cross-Industry Remote Condition Monitoring Strategy Group established a Cross-Industry RCM (XIRCM) research program to tackle this issue, namely the T1010 project [3]. The RSSB and Network Rail presented the first results of this research project at the IET RCM conference in 2014 [4]. To properly generate business cases for any new monitoring sensing hardware, it is essential to assign value to the data generated by one party but used by another. To tackle the cost issue, it was proposed in project T1010 that a commercial agreement should be made between the parties who might be involved before installing a new monitoring sensing system [5]. The commercial agreement has limitations that hinder complete data management and cost attribution between

stakeholders. It does not solve the need for a trusted third party to enforce conformance to the agreement. In addition, it does not provide any evidence for the absence of misbehavior of the parties involved. Therefore, we believe that leveraging technology such as blockchain will have more impact on cost attribution and data delivery between parties and could solve the aforementioned problems efficiently. The auditable and secure nature of blockchain technology will encourage stakeholders to share RCM data and share the cost of their data in a fair way. As blockchain embodies three main protocols, decentralization, cryptography, and consensus, adoption of this technology will promote interoperability among business processes and stakeholders [6].

The structure of this paper is as follows: in section II, we will give a brief background about use of RCM data and blockchain in the rail industry. In section III, related work will be investigated and discussed. In section IV, our framework will be described in detail. In sections V and VI, implementation and future work, and a conclusion will be presented, respectively.

II. BACKGROUND

Blockchain-driven technologies help to improve transactions and make them more efficient and secure due to the censorship-resistant and tamper-proof digital platforms of distributed trust provided by this disruptive technology. Blockchain is still in its infancy and it needs more work to reach maturity. However, as the technology ecosystem continues to mature, considerable efforts have been made to explore its applicability and potential diffusion in different sectors including the industrial sector [7], [8]. The main barriers to adopting this disruptive technology in industry have been investigated and analyzed [9]. In the rail industry, blockchain-based systems have already been implemented for ticket sales, invoicing, and freight consignment, among others [10].

In the UK, a large volume of data is generated daily from RCM, which is attracting considerable attention as it holds great value to enhance operation and maintenance of the whole system. Thus, this is an area of active research and several projects have been initiated focusing on this area [11]. To date, railway parties get the information they need by actively searching for relevant data in different places. In spite of the availability of different related data sources, accurate delivery of relevant, timely information to these railway parties is still inadequate. The multi-party and sensor ecosystem of the railway industry means that RCM of the railway network has become an "Internet of Railway Things" (IoRT) [12]. Thus, integration of state-of-the-art IT, the IoT, cloud computing, and big data is creating the feasibility of "smart railways" [13].

The data generated in RCM takes several forms depending on the sensing source of that data, including audio, video, pictorial, continuous analogue measurements, and digital threshold signals. The raw data produced will be enhanced and processed to be used in future analytics to manage the asset health and lifetime. In condition monitoring, there are six

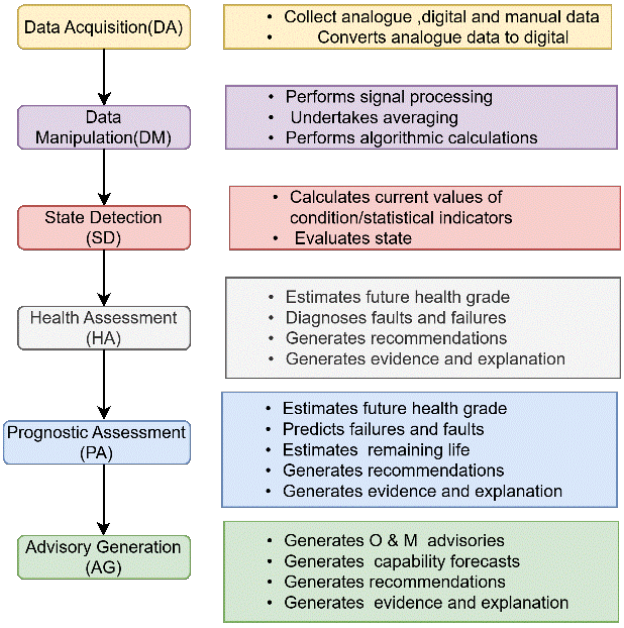

Fig. 1. The Six Processing Levels of ISO 13374. Source: [14]

recognized levels of data processing as illustrated in Fig.1 [14], ranging from simple data collection, through the sending of alarms when a certain condition is reached, to a complete diagnostic capability which includes sending notifications to the operations and maintenance team to instruct them to fix a specific asset before it fails.

The source of each data level might belong to different stakeholders and the value of data becomes greater when the level of data processing is higher. According to [5], unless contractual clauses have stated the opposite, the Intellectual Property Rights (IPR) of data recorded by specific equipment are vested in the party that produced the data, while the IPR of data resulting from applying an improvement or modification on others' data shall be vested in the party who has made the improvement or modification. To share the data they own, owners expect some kind of compensation. As such, the value of data may be evaluated and a "price" set by the owner to be paid by a consumer for access to the data.

We have to keep in consideration, similar to any trading environment, that both the data provider and the consumer need to build a trusted relationship through the trading system. To this end, the system is expected to enforce that both provider and consumer conform to the agreement and assure the accuracy of the data delivered. In addition, it has to provide a traceable way to allocate any occurrence of data tampering to be used as evidence that affects processing payment, compensation, or reimbursement.

In current situations, when one party refrains from making the agreed payment or service whether consumer or provider, a Trusted Third Party (TTP) such as a bank, third escrow mediator, or dispute board might be a necessity which creates a kind of bureaucracy and introduces additional costs. Thus, to

dispense with the need for a TTP and to protect the monitoring data provided from being tampered with, blockchain and smart contracts might be used as they provide immutable data storage and a supporting trusted payment system. We can use Smart Contracts (SC) to define all the quantitative and qualitative terms between providers and consumers in a tamper-proof manner. An SC is defined as an executable script deployed to run in the blockchain in a distributed manner [15].

Exploiting the characteristics of blockchain to overcome the abovementioned threats is our goal in this research. Data immutability, auditability, decentralization, and the emerging concept of SC will make it possible to build solutions, taking RCM in the rail industry to very advanced level. As pointed out by Christidis and Devetsikiotis [16], an SC can be seen as a stored procedure in a relational database management system. By using SCs, a wide range of applications can be developed based on the underlying execution platform provided by the blockchain.

In this research, we present an initial design of a blockchain-based framework which simplifies and automates the cost attribution process between providers and consumers. By employing SCs, the provider and consumer will be assured that their agreement has been enforced and the terms will not be changed. Moreover, the TTP can be dispensed with as the SC will work as the trusted mediator.

We can summarize the contribution in the following:

- Translation of agreement terms into blockchain-based SCs
- Automated and fair cost attribution and service fees
- Transparent and immutable agreements and cost attribution
- Data IPR processing service
- Integrity-proof data system

III. Related work

To the best of our knowledge, the proposed blockchain-based framework has not been proposed before in the rail industry or any other industrial area, but research on the monetization of IoT data based on blockchain technology has gained in intensity in the few last years. Many have proposed SC-based schemes for trading and sharing IoT data over a blockchain network, with models differing from our proposal totally or partially. In BlockSubPay [17], Oktian et al proposed a payment protocol using the Ethereum blockchain network. In their proposal, the network will record subscriptions related to each user, which could be fixed payment or pay-as-you-go subscriptions. Initiation of the subscription process will be off-chain and the cloud provider will provide the client with the address of the SC that complies with the client's preferred payment method. The subscriber will hold a token that is used to access the provider's resources on the cloud. The client has to manage the access tokens they might have if subscribed to several data providers. In Saranyu [18], Nayak et al proposed a cloud tenant and service management system similar to [17], but they used Quorum as the platform, a permissioned network to implement SCs, and have not provided enough details about charging tenants. Al-Zahrani [19] also proposed a subscription-based model for trading cloud services data. In his proposed model, all subscription requests are recorded to the ledger, even those for which the payment has not been completed. In [20], a blockchain-based solution was implemented using Ethereum to automate both payment and the issue of random tokens to the IoT owners. The data user makes an ether deposit when subscribing to a specific IoT device and before accessing the data which is stored in the MQTT broker which represents the central point of failure. Excluding [18], the aforementioned works have not provided a solution to suspend or revoke the subscription of malicious users other than removing the data from the cloud resource. Moreover, in all the mentioned works, the authors assumed an honest provider and did not discuss the presence of falsified or garbage data to mislead consumers. The double deposit escrow has been proposed before in online BitPay [21], BitHalo [22], and DCSP [23]. In all three, both client and provider build the escrow through an SC for deposit values only, but the main payment is done off-chain. To release the escrow, both parties must confirm the transaction is successful. Otherwise, they will lose their deposit as they have not provided any dispute solutions. Asgaonkar and Krishnamachari [24] proposed a dual-deposit escrow scheme similar to the previous three schemes but which does provide a dispute solution and include the main payment. Still, their scheme serves only for one-time access and the buyer has to evaluate the transaction and send a response to unlock the escrow and process the payment. If the buyer does not respond, or was misbehaving, the seller never gets compensation for that and may lose their deposit and right to the payment. In [25], a different framework for data sharing is proposed in which the hash values of data are encrypted with a symmetrical key and stored off-chain on the cloud by the provider ahead of the sharing process. All providers will advertise their data offers and public keys on the cloud. The SC is generated on fly to provide only one-time access to the data and the generated agreement will be stored on-chain to be used for future dispute solutions. The solution they provide to face any breach through a voting process is not detailed enough in the description and implementation sections. In our model, the escrow is released based on the agreement status and our dispute solutions will assure avoidance of escrow locking or loss of payment/compensation. The blockchain trilemma [26] refers to the scalability, decentralization, and security characteristics which cannot be fulfilled simultaneously. Therefore, all the works discussed above chose to store a large volume of data off-chain to avoid any scalability issue. Similarly, scalability has major implications in blockchain applications in RCM due to the huge volume of data generated of diverse types. To solve this issue, hash values will be stored on-chain since they are used as proof of data ownership and will be used to automate integrity and verify latency claims.

IV. The proposed framework

In this section, the proposed approach is described, and all introductory information, assumptions, and decisions backing

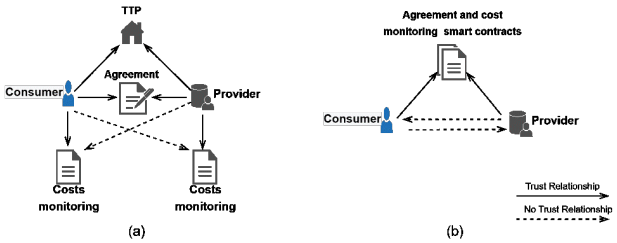

Fig. 2. Trust Relationship between actors

the framework are illustrated. To inform the design of the blockchain solution, the trust relationships between actors need to be identified.

The actors in our proposed future process will be the data owner (provider), data user (consumer), and blockchain-hosted SCs, removing the need for the TTP as an actor involved in the current process.

Fig.2(a) depicts the trust relationship among the actors who take part in the current cost attribution and payment process. In this situation, both have to trust that the other party will comply with the agreement that specifies the cost attribution, the TTP, and their respective assessment of local financial cost, henceforth referred to as "local cost monitoring". Both provider and consumer have full control of their local cost monitoring, and this might lead to a dispute as it is not protected against tampering or misunderstanding. Each party might have a different or dishonest evaluation for each part of the signed agreement in terms of service quality. Providers may pretend to comply with the signed agreement and provide the demanded service quality they advertised at the beginning. Based on that, they would expect a cost attribution based on their own evaluation. On the other side, consumers may pay more for lower service quality or may get fluctuating quality; this will affect their evaluation of the cost they should pay. Both parties may provide evidence to strengthen their point of view but as there is no trust between them, there will be no trust in the correctness of their evidence. In addition, the involvement of a TTP in processing payments, processing IPR, or solving disputes, will burden the system with all the consequences of the bureaucracy and its additional costs in terms of money and time.

Fig.2 (b), depicts the trust relationship in our proposal. There is no need for a direct trust relationship between provider and consumer. Both parties focus their trust on the SCs.

A. Design Decision and Assumptions

Our proposal will be implemented using a permissioned blockchain network. A trusted authority is needed to authorize the participation node in the blockchain network. In our proposal, the Department for Transport (DfT) is our candidate for this role as it is supposed to be the most neutral party that dominates the strategic framework for transport services. Thus, data providers and data consumers will be known to each other and hold unique membership identities over the blockchain network. Establishment of identity relies on the existing identity manager that is often a part of blockchain implementation such as the Membership Service Provider (MSP) and Certificate Authority (CA) in Hyperledger Fabric [29]. To maintain scalability, off-chain storage is used, in which some data can be encrypted and moved to be stored outside the blockchain. At the same time, references to this data are stored on-chain and used to confirm the correctness of the off-chain source data. The SCs, which have the capability to manage and monitor data transmissions, are able to confirm access eligibility and data correctness (consumption) in disputation cases, without recording the actual data on the blockchain. Since the on-chain data will be available to all parties that are granted network permissions, confidentiality of sensitive data such as pricing and cost can still be assured. This is achieved by adopting encryption methods to protect such sensitive data from being accessed by parties not directly involved in a specific data access agreement.

B. Interaction of Actors in the Network

Fig.3, depicts the main interactions between actors in the blockchain network.

- A **provider** who has data to trade needs to create an offer and push it to the network. After accepting the offer, the provider will be able to upload the hash values of their data *h(data)* to the blockchain. This hashed data will serve as a fingerprint to connect this provider to the generated data as an owner. It can be used later if the consumer complains about data integrity or if the provider complains that other providers are reselling data and breaching their IPR. It is the provider's responsibility to upload and update the hash values regularly to prove their rights to the generated data and to provide proof of data integrity to consumers. Offer availability may change when the provider is no longer able to offer the data, by changing the *Validity* attribute to False. As a result, the provider will not be allowed to upload new hash values unless they have an ongoing offer. In addition, all ongoing agreements will be revoked and finalized.
  The provider will encrypt the original data which will be stored off-chain using the consumer's public key after signing the data using the provider's private key. This will secure the data if any disclosure has occurred. The signature is important for use later if the consumer complains about data integrity or data corruption. To this end, the provider's signature will be examined to prove that they are the source of the data. In the rail industry, a provider could be any stakeholder who funds or operates sensors for RCM and retains ownership of the generated data.
- A **consumer** will be able to list all registered offers and, when interested in a specific offer, can send a request over the network. The consumer's request will be accompanied by the subscription period to initiate a new payment process then a new possible access agreement. By creating an access agreement, the consumer will be able to access

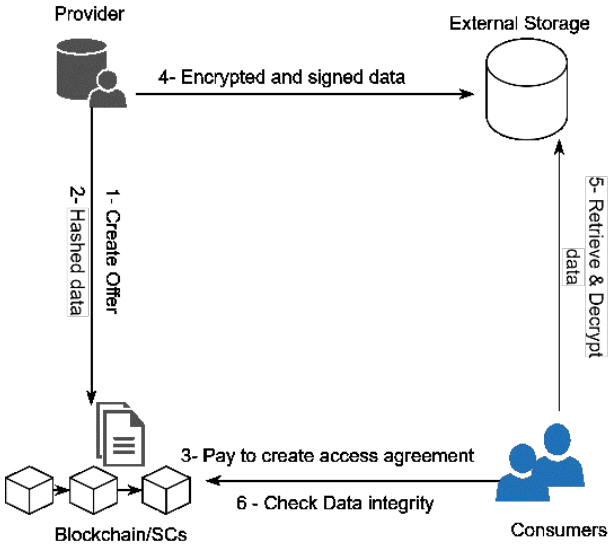

Fig. 3. main interactions between actors.

the hash values uploaded by the provider to the network during the period of their ongoing agreement. If the agreement is terminated due to expiration or revocation, their access to the hash values will be terminated as well. In the rail industry, a consumer could be any authorized stakeholder who needs data to manage and maintain asset condition.

- **SCs** are based on the accounting and data access models adopted, and are expected to monitor cost calculations, data delivery, and an automatic payment process without the need for a TTP. These processes are described in detail in the following sections.

C. Data Access Agreement Model

The commercial agreements described in project T1010 [5] have guided our selection of the components included in the SC regarding agreement between the data provider and data consumer

Fig.4 illustrates the structures of these components and for the sake of simplicity we have chosen the most relevant ones but of course this can be modified and expanded in the future. For each new agreement between provider and consumer, two new records will be added to the ledgers that are shared to them. The first one will hold information on the settled agreement between the data consumer and data provider. The second one is an accounting record used in enforcing the data cost and compensation. To assure the data IPR, the data provider is the one who has the ownership of the provided data. Thus, no one else has the right to advertise an offer for the same data and this is warranted by the uploaded hash values of the data. The data consumer could be a data provider after developing and modifying the procured data to generate data with a higher level of processing before advertising data to the network for others. This process will lead to the generation of different hash values for the modified data which will connect

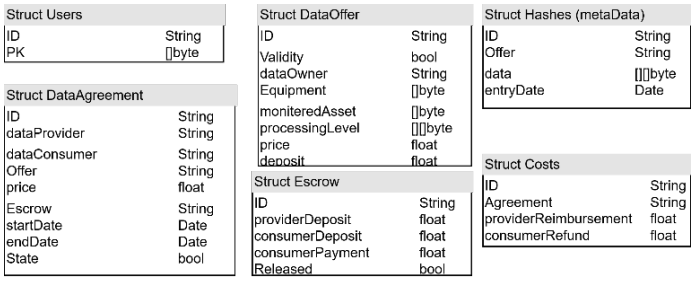

Fig. 4. Data structure.

it to the new owner and keep their right to create a new offer to advertise their data.

In the setting-up phase of the system, all nodes of either providers or consumers must be registered with the trusted authority and they must have their IDs and public/private key pair before participating. The adopted consensus mechanism should be chosen as well.

The process flow is as follows :

1) The consumer will send a request determining the offer (offer ID) they are interested in to the SC along with the subscription period and payment
2) The SC will check the validity of the requested offer. If not valid, then the request will be rejected. If the offer is still available, a payment process is initiated; this process is discussed in detail in the following section.
3) If the payment process is completed, an agreement between the provider and consumer will be generated automatically.
4) Both provider and consumer will be notified about the established agreement.
5) The provider will use the consumer's public key to encrypt the original data and their private key to sign it before uploading the data onto the external storage.
6) Based on the agreement, the consumer will gain access to the original data on an off-chain channel (external storage) and access to the hash values on-chain. Original data should be signed then encrypted on the external storage as follows: $consumer_{PublicKey}(provider_{PrivateKey}(Data))$.
7) The consumer will decrypt the data and hash it to compare it with the hash values provided on-chain to check its integrity.
8) It is the consumer's responsibility to use the hash values to ensure the integrity of the original data. Two types of malicious behavior from the provider side can be proven in our model:
   a) Sending corrupted or incomplete data;
   b) Latency in providing hash values to the consumer.
9) The agreement generated can be revoked before the dedicated expiry date by the provider or the consumer. This process is irreversible, i.e. when the state is changed to False (revoked), no one is allowed to activate it again to True. Instead, a new agreement must be initiated from

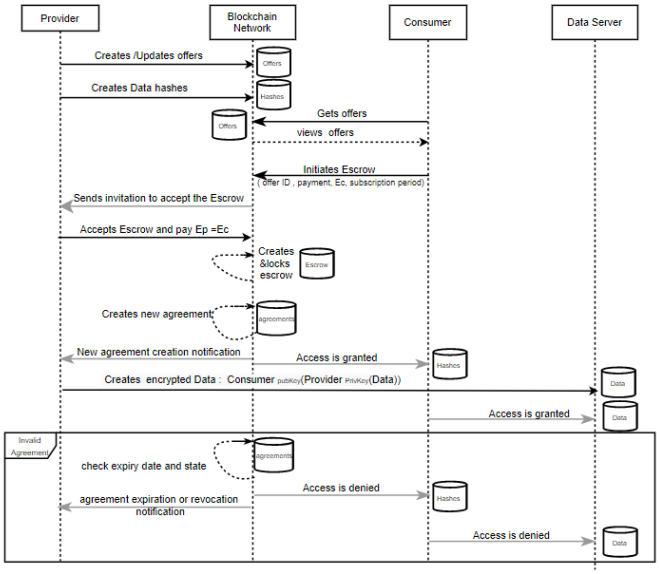

Fig. 5. Data access agreement sequence model.

step 1.

Fig.5 illustrates the sequence of creating the data access agreement.

D. Accounting Model

In any trading platform, the payment process can be managed using a post-paid or pre-paid model. The former implies a trust in the consumer (buyer) that if the data is received correctly first, the payment will be made as agreed. The latter implies a trust in the provider that if the payment is made as agreed, the data will be sent after. None of the previous models guarantee full satisfaction for both provider and consumer and both carry some risk in case one party decides to misbehave. Therefore, there is a need for a TTP to provide an escrow service for both provider and consumer.

In our proposal, to eliminate the need for a TTP, an SC will be used as an escrow which holds the consumer's payment and performs the process of payment to the provider after the data is sent and the consumer confirms it has been received as agreed, which will be implied by the consumer continuing their agreement with the provider without revoking the request. Also, the escrow SC will hold the penalties that both provider and consumer would pay in advance of any data transfer process. Penalties will be used if there is malicious behavior by either or both. Otherwise, the penalties are reimbursed to both parties if there is no complaint and the data trading is going smoothly conforming to the agreement.

With each offer, the provider is responsible for deploying the following attributes and values as depicted in Fig.4:

D_p: The price of the data to be shared in a determined period (daily/weekly/monthly/quarterly/annually).

E : The deposit both consumer (E_c) and provider (E_p) should pay to build an escrow as a penalty for any behavior not conforming to the agreement. This value may be determined by a special process by following predefined legalization rules. The determination of prices and deposit values is an economic topic in the rail industry that goes far beyond our aim in this work. Since it is out of our scope in this phase, we will assume that accepted offers have applied the predefined restrictions and rules.

$h(D)$: The hash value of the data this provider is sharing. Payment process flow is as follows:

1) When the consumer selects an offer, they will initiate the escrow SC by sending the payment and the deposit as determined in the offer.
2) The provider will be notified of the request and will check the payment and deposit; if they both match their agreement, the provider will pay their deposit, which should not be less than the consumer's deposit, to lock the escrow. Otherwise, if the payment does not match, the process will be rejected and the consumer will get their payment back.
3) When the three values (payment, consumer's deposit, and provider's deposit) are paid, an SC to initiate an agreement will be triggered, in which the period of access to the data is determined.
4) Based on the consumer's satisfaction, they will decide to proceed with the agreement or revoke it, or claim a breach from the provider's side by providing the $(provider_{privateKey}(Data))$ at any point. The provider also has the same ability to revoke the agreement but they would lose their deposit in that case.
   - If the agreement period has ended without any revocation or claim, both will receive their deposit back without any deductions. In addition, the provider will receive the payment as agreed. Similarly, In the case of revocation without any claim, the payment will be calculated based on the period the consumer had access to the data before revoking the agreement;
   - If the consumer submits a claim that is one of two kinds, a mismatched claim or latency claim, the following applies:

     a) In a mismatch claim, the consumer will provide the received data which is signed using the provider's private key to the SC alongside the date of receipt. Then, the SC will hash the provided data after verifying the source using the provider's public key. Next, the hashed value will be compared to the one stored on-chain. In the case of a mismatch, the consumer will be refunded part of the payment based on the claim date and will receive both deposits as compensation. The agreement between the consumer and provider will be revoked at this point to prevent the consumer from future free access. In contrast, if the hashes match, the provider will receive payment based on the claim date and will receive both deposits as a penalty for this frivolous complaint. The agreement between them will be

revoked as well.

b) In a latency claim, the dates of hashed data on the chain will be checked to verify whether or not the provider submitted the hashed data on an acceptable timeline using the block's timestamp. Also, the agreement will be revoked.

Fig.6 shows a graph tree depicting all the possible scenarios in trading data between the consumer and provider and how the cost is calculated based on the behavior of consumer and provider.

Before illustrating how the cost will be calculated in each scenario, we will define the acronyms used in the following equations:

- $C_{Payment}$:The initial payment the consumer will make to initiate the escrow along with E_c. This payment represents the final price of the whole life of the agreement.
- Act_{Price} : The actual price of the consumed data based on the service period; this value should be less than or equal to $C_{Payment}$.
- $P_{Reimbursement}$: The actual reimbursement that will be transferred to the provider depending on the service period and consumer satisfaction, which is reflected by agreement status and resolution of any claims.
- C_{Refund}: Refunds which will be transferred to the consumer based on the agreement status and resolution of any claims.

Three different dates will affect the actual price value (Act_{Price}):

- Rev_{Date} : The revocation date which will be the same as the time stamp of the block including the revoked agreement.
- $Start_{Date}$: The agreement's start date; this will be declared in the agreement as an attribute.
- End_{Date}: The agreement's end date; this will be declared in the agreement as an attribute.

Scenario A: The provider sends genuine data and the consumer has a real claim about the latency in appending the hashes to the network. The agreement will be revoked, triggering the SC to calculate the costs as follows:

$$\begin{aligned} Act_{Price} &= D_p * (Rev_{Date} – Start_{Date}) \\ P_{Reimbursement} &= Act_{Price} \\ C_{Refund} &= (C_{Payment} – Act_{Price}) + E_P + E_C \end{aligned} \tag{1}$$

Scenario B: The provider sends genuine data and the consumer makes a frivolous claim about data integrity or latency in appending the hashes to the network. The agreement will be revoked, triggering the SC to calculate the costs as follows:

$$\begin{aligned} Act_{Price} &= D_p * (Rev_{Date} – Start_{Date}) \\ P_{Reimbursement} &= Act_{Price} + E_P + E_C \\ C_{Refund} &= C_{Payment} – Act_{Price} \end{aligned} \tag{2}$$

Scenario C: The provider sends genuine data and the consumer makes no claim but wants to revoke the agreement for another reason such as finding another resource for the same service. The agreement will be revoked, triggering the SC to calculate the costs as follows:

$$\begin{aligned} Act_{Price} &= D_p * (Rev_{Date} – Start_{Date}) \\ P_{Reimbursement} &= Act_{Price} + E_P \\ C_{Refund} &= (C_{Payment} – Act_{Price}) + E_C \end{aligned} \tag{3}$$

A similar scenario occurs when the agreement expires without any revocation or complaint from the consumer's side:

$$\begin{aligned} Act_{Price} &= D_p * (End_{Date} – Start_{Date}) \\ P_{Reimbursement} &= Act_{Price} + E_P \\ C_{Refund} &= (C_{Payment} – Act_{Price}) + E_C \end{aligned} \tag{4}$$

Scenario D: The provider sends falsified data and the consumer makes a real claim about data integrity or latency in appending the hashes to the network if was not aware about the falsified data. The agreement will be revoked, triggering the SC to calculate costs in a similar way to equation (1) in scenario A. The consumer will be compensated and the provider will get part of the total payment and will lose their deposit.

Scenario E: The provider sends falsified data and the consumer makes a frivolous claim about the latency in appending the hashes to the network. The agreement will be revoked, triggering the SC to calculate the costs in similar way to equation (2) in scenario B. The provider will be compensated and the consumer will lose their deposit.

Unfortunately, as the consumer is the only source of the actual received data, if they do not recognize falsified data, they will lose their deposit and the costs calculation will be the same as in scenario B. As everything is recorded permanently on the network, the consumer can raise a claim to the dispute board and prove it at any time after the agreement is revoked as long as they still have the signed original data.

Scenario F: The provider chooses to revoke an agreement as they are no longer capable of providing the data. Costs will be calculated in similar way to equation (1) in scenario A by which the consumer will be compensated and the provider will get part of the total payment and will lose their deposit.

In our proposal, the malicious behaviors that can be proven against the consumer are frivolous complaints and data reselling. If the provider revokes the agreement due to notifying provable malicious behavior such as data reselling, they can raise the claim with the proof; otherwise, they will lose their deposit. Actually, to prove data reselling, the provider should get access to the hash values as a consumer first. A provider is not allowed access to the hash values appended by other providers unless they are one of the consumers. This type of case may lead to legal action and should be managed by the dispute board.

V. IMPLEMENTATION AND FUTURE WORK

Based on their permission levels, blockchains are primarily categorized as public/permissionless blockchain networks or private/permissioned blockchain networks. A public blockchain network, e.g. Bitcoin, is open to the public to

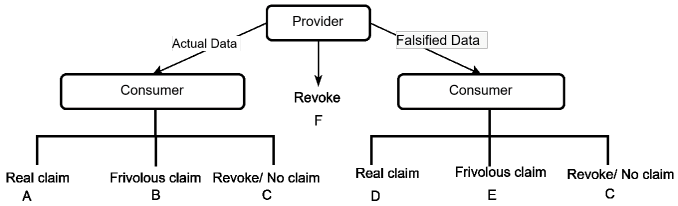

Fig. 6. All possible scenarios in trading data.

join, i.e. anonymous participants have uniform access privileges to the network ledger. Therefore, it imposes a powerful consensus mechanism to preserve security; the most popular consensus algorithm used in public blockchains is Proof-of-Work [27]. In contrast, in private networks, the participant should be known and identified to restrict their access to the ledger. In our framework, the identities of all the participants in the network should be known. Therefore, the proof of concept will be implemented using a blockchain network supporting the auditing and tracking process. There are several blockchain platforms that could be employed in implementing the proposed model. Determining the most appropriate one has a considerable influence on the design as there are no one-size-fits-all-platform blockchain initiatives. Therefore, a tradeoff analysis was conducted on the most used blockchain platforms: Ethereum [27], [28], Fabric [29], Sawtooth [30], and Iroha [31] based on the criteria listed below in Table I.

TABLE I
TRADEOFF ANALYSIS BETWEEN ETHEREUM, FABRIC, SAWTOOTH, AND IROHA.

Criteria	Blockchain Platforms			
	Ethereum	*Fabric*	*Sawtooth*	*Iroha*
Supports smart contracts.	✓	✓	✓	✓
Consensus algorithm modularity.	×	✓	✓	×
Built-in components for managing. identities.	×	✓	×	✓
Supports payment in fiat currency.	×	✓	✓	✓
Proficient in maintaining different privacy levels between users.	×	✓	✓	✓

Automation of all the processes discussed will be through an SC that is implemented in Turing-complete languages. Ethereum uses Solidity, a new programing language that provides reasonable but is expensive and to some extent limited in implementing complex contract terms. The other platforms support more profound programing languages such as Java, Go, Rust, and C++. Additionally, Iroha is focused heavily on supporting the development of mobile applications and embedded systems alongside web applications. This platform provides a set of libraries and prebuilt components including predefined SCs and queries that will facilitate the adoption of distributed ledger technologies into IoT infrastructure. Therefore, Iroha might be useful in complementing Fabric and Sawtooth platforms by providing reusable components.

Ethereum, as a public blockchain in origin, handles the abuse of trust by imposing a proof-of-work (PoW) consensus algorithm which is known to be rigorous but power and time consuming due to the mining process and propagation, while the Fabric and Sawtooth platforms supports several consensus algorithms which can be changed on the fly while the network is running, which make them adaptive to different environments. Additionally, consensus modularity will give us the ability to implement our proposal with different consensus mechanisms to examine and measure the throughputs according to each one. Iroha embraces its own consensus algorithm, a crash fault-tolerant consensus called Yet Another (aka YAC). In fact, consensus protocols in private blockchain avoid all unnecessary hurdles and complexities since reaching a total agreement on the common truth between predefined identities will be easier and faster.

Ethereum maintains anonymity in a way that any node can join or leave with no restrictions. This does not serve our purposes in this research because it is essential to identify each participant in the network. Hyperledger Fabric provides MSP and CA services to identify the participant in an easy and manageable way. Sawtooth does not have a CA service similar to the one in Fabric, thus the developer might need to integrate external identity software. Iroha, on the other hand, has an intrinsic support for identity management as well.

Ethereum incurs fees (gas) in exchange for every SC execution and has its own native payment currency (Ether) while the Hyperledger platforms Fabric, Sawtooth, and Iroha are cryptocurrency-independent and payment in fiat currencies is available.

The proposed framework dictates the variation of privacy level between users, i.e. not all agreements and payment processes are available for all network users. Some users may choose to have a private agreement and keep the cost attribution hidden from others who are not involved in that agreement. Ethereum maintains an identical role for all participants, and all transactions are available and visible to all participants in the network, while Hyperledger platforms are able to satisfy this requirement by one way or another. In Fabric, this issue is managed by creating a separate channel to isolate participants that need private agreements and cost attributions, while in Sawtooth, changing the identity namespace in the transaction family will restrict access to certain identities. Similarly, in Iroha, defining access control rules will maintain convenient role-based access at different levels.

All in all, a permissioned blockchain network seems to be the best choice to fulfill the predefined design decisions we mentioned before in our framework when considering faster settlement, scalable performance, and a more controlled environment.

In order to evaluate our proposal to improve trust, automate a fair cost attribution process and payment, and enforce agreements between parties, we look to apply it to related cases in the rail industry. To this end, two case studies will be examined against the proposed framework in our empirical study to validate the outcome. The first case study will be the Unattended Overhead Line Equipment Monitoring System (UOMS), a train-based system monitoring infrastructure. The

second will be an acoustic axle bearing monitoring system (RailBAM), an infrastructure-based system monitoring trains. In UOMS, equipment is mounted on a CL390 train and used to monitor and measure the health of the pantograph line which belongs to the infrastructure. In the second case study, RailBAM, the acoustic devices are mounted on the main infrastructure track and used to monitor the axle journal bearing upon which the wheel of rolling stock is rotating. Both case studies involve cooperation between different stakeholders across the rail industry. In other words, there are several parties in the rail industry interested in the data generated in both case studies, such as Network Rail, Train Operating Companies (TOCs), Freight Operating Companies (FOCs), and other train manufacturers and maintainers. The results of applying our proposal using these two case studies will be reported to the community in future.

VI. CONCLUSION

Cross-industry RCM is a vital process in the rail industry owing to its necessity to improve service quality and safety, and to save costs. Therefore, intelligent devices and sensors are extensively attached to infrastructure and trains to capture data about rail assets and then used in decision-making and maintenance process guidance. To derive the maximum benefits, it is essential to coordinate the business process among the stakeholders who pay for these devices and those who gain access to the data to obtain knowledge from this collected data about their assets. In fact, trading data is providing a net business benefit as this data will eventually be used to improve railway service quality and cut costs.

The current process adopted to regulate agreements between facilitators and beneficiaries is not settled well and is done in very conventional way. Moreover, some of the coordinating agreements are gentlemen's agreements that are easy to tamper with and create untraceable dispute evidence. That is, they may lead to unfair cost attribution, disputes between parties, and misuse in some cases. In addition, the current agreements do not protect the facilitator's rights in owning the data as they are not tamper-resistant.

We have introduced a new framework based on blockchain technology to solve all the aforementioned issues. Our framework regulates the whole process and grants data ownership to the facilitator as long as they provide the hash values to the blockchain network. It simplifies the communication process between data provider and data consumers, automates agreement building, and automates cost attribution based on the built agreement. To some extent, the proposed framework provides a service level agreement between provider and consumer, i.e. both may claim some breaching behaviors. For example, a consumer may claim poor service quality, prove their claim, and be compensated; otherwise, the consumer will be fined for making a frivolous claim.

We are hoping the proposed framework is tailored in a way that will solve most of the vulnerabilities in the current situation and will move the condition monitoring process to a higher maturity level. We believe it will encourage all stakeholders to engage in protecting their data and to generate some revenue from their data. All the anticipated features will be examined in the empirical study and we will report to the community in the near future.

ACKNOWLEDGMENT

This project has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 826156. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union. The authors would further like to acknowledge Imam Abdulrahman Bin Faisal University and the Saudi Government for funding R.A.A., the first author.

REFERENCES

[1] C. P. Ward et al, "Condition monitoring opportunities using vehicle-based sensors," Proceedings of the Institution of Mechanical Engineers. Part F, Journal of Rail and Rapid Transit, vol. 225, (2), pp. 202-218, 2011.

[2] A. Alemi, F. Corman, and G. Lodewijks, "Condition monitoring approaches for the detection of railway wheel defects," Proceedings of the Institution of Mechanical Engineers. Part F, Journal of Rail and Rapid Transit, vol. 231, (8), pp. 961-981, 2017.

[3] Sparkrail, Cross-industry remote condition monitoring (T1010). [Online]. Available: http://www.sparkrail.org/Lists/Records/DispForm.aspx?ID=8096

[4] G. J. Tucker and A. Hall, "Breaking down the barriers to more cross-industry Remote Condition Monitoring (RCM)," in 6th IET Conference on Railway Condition Monitoring (RCM 2014), Birmingham, UK, September 2014, pp. 1-6.

[5] Sparkrail, Cross-industry remote condition monitoring, Commercial, Final report Appendix E Standard Form (Template) (T1010 Report Appendix).

[6] M. Crosby, P. Pattanayak, S. Verma, and V. Kalyanaraman, "Blockchain technology: Beyond bitcoin," Applied Innovation, vol. 2, pp. 6-10, 2016.

[7] M. Friedlmaier, A. Tumasjan, and I. M. Welpe, "Disrupting industries with blockchain: The industry, venture capital funding, and regional distribution of blockchain ventures," in Proceedings of the 51st Hawaii International Conference on System Sciences, 2016.

[8] M. Risius and K. Spohrer, "A blockchain research framework," Business & Information Systems Engineering, vol. 59, (6), pp. 385-409, 2017.

[9] B. Biswas and R. Gupta, "Analysis of barriers to implement blockchain in industry and service sectors," Computers & Industrial Engineering, vol. 136, pp. 225-241, 2019.

[10] P. Mcmahon, T. Zhang, and R. Dwight, "Requirements for big data adoption for railway asset management," IEEE Access, vol. 8, pp. 15543-15564, 2020.

[11] D. Galar, D. Seneviratne, and U. Kumar, "Big data in railway O&M: A dependability approach," in S. Kohli, A. V. Senthil Kumar, J. M. Easton, and C. Roberts (Eds.), Innovative applications of big data in the railway industry. IGI Global, Hershey, PA, pp. 1-26, 2017.

[12] J. M. Easton. "Blockchains: A distributed data ledger for the rail industry," in S. Kohli, A. V. Senthil Kumar, J. M. Easton, and C. Roberts (Eds.), Innovative applications of big data in the railway industry. IGI Global, Hershey, PA, pp. 27-39, 2017.

[13] Q. Y. Li et al, "Chapter 14 - Smart railway based on the Internet of Things," in H.-H. Hsu, C.-Y. Chang, and C.-H. Hsu (Eds.), Big data analytics for sensor-network collected intelligence. Elsevier Inc., pp. 280-297, 2017.

[14] ISO13374-2: "Condition monitoring and diagnostics of machines – Data processing, communication and presentation – Part 2: Data processing", 2007.

[15] M. Alharby, A. Aldweesh, and A. V. Moorsel, "Blockchain-based smart contracts: A systematic mapping study of academic research," in International Conference on Cloud Computing, Big Data and Blockchain (ICCBB), Fuzhou, China, 2018, pp. 1-6.

[16] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," IEEE Access, vol. 4, pp. 2292-2303, 2016.

[17] Y. E. Oktian, E. N. Witanto, S. Kumi, and S. Lee, "BlockSubPay - A blockchain framework for subscription-based payment in cloud service," in 2019 21st International Conference on Advanced Communication Technology (ICACT), PyeongChang Kwangwoon_Do, Korea (South), 2019, pp. 153-158.
[18] S. Nayak, N. C. Narendra, A. Shukla, and J. Kempf, "Saranyu: Using smart contracts and blockchain for cloud tenant management," in 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), San Francisco, CA, 2018, pp. 857-861.
[19] F. A. Al-Zahrani, "Subscription-based data-sharing model using blockchain and data as a service," IEEE Access, vol. 8, pp. 115966-115981, 2020.
[20] A. Suliman, Z. Husain, M. Abououf, M. Alblooshi, and K. Salah, "Monetization of IoT data using smart contracts," IET Networks, vol. 8, pp. 32-37, January 2019.
[21] Bit-Bay, Double deposit escrow. [Online]. Available: https://bitbay.market/double-deposit-escrow
[22] D. Zimbeck, Two party double deposit trustless escrow in cryptographic networks and bitcoin. [Online], 2014. Available: https://bithalo.org/whitepaper_twosided.pdf
[23] G. Bigi, A. Bracciali, G. Meacci, and E. Tuosto, "Validation of decentralised smart contracts through game theory and formal methods," in C. Bodei, G. Ferrari, and C. Priami (Eds.), Programming languages with applications to biology and security. Springer, pp. 142-161, 2015.
[24] A. Asgaonkar and B. Krishnamachari, "Solving the buyer and seller's dilemma: A dual-deposit escrow smart contract for provably cheat-proof delivery and payment for a digital good without a trusted mediator," in 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Seoul, Korea (South), 2019, pp. 262-267.
[25] H. Desai, K. Liu, M. Kantarcioglu, and L. Kagal, "Adjudicating violations in data sharing agreements using smart contracts," in 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Halifax, NS, Canada, 2018, pp. 1553-1560.
[26] T. Ometoruwa, Solving the blockchain trilemma: Decentralization, security & scalability. [Online]. Available: https://www.coinbureau.com/analysis/solving-blockchain-trilemma
[27] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," Technical report. [Online]. Available: https://gavwood.com/paper.pdf
[28] V. Buterin, Ethereum white-paper. [Online]. Available: https://ethereum.org/en/whitepaper/
[29] E. Androulaki et al, "Hyperledger fabric: A distributed operating system for permissioned blockchains," in EuroSys '18: Proceedings of the Thirteenth EuroSys Conference, April 2018, pp. 1-15.
[30] K. Olson et al, Sawtooth: An introduction – White paper. [Online]. Available: https://www. hyperledger.org/wp-content/uploads/2018/01/Hyperledger_Sawtooth_WhitePaper.pdf
[31] Hyperledger Iroha Community, Iroha handbook: Installation, getting started, API, guides, and troubleshooting. [Online]. Available: https://iroha.readthedocs.io/_/downloads/en/1.1.3/pdf/